Informational Cascades and Software Adoption on the Internet: An Empirical Investigation

Article *in* MIS Quarterly · March 2009

DOI: 10.2307/20650277 · Source: DBLP

3 authors, including:

Bin Gu
Boston University
104 PUBLICATIONS **4,835** CITATIONS

Andrew Whinston
University of Texas at Austin
686 PUBLICATIONS **18,897** CITATIONS

Informational Cascades and Software Adoption on the Internet:

An Empirical Investigation

Wenjing Duan

Funger 515, School of Business

The George Washington University, Washington, DC 20052

Phone: (202) 994-3217 Fax: (202) 994-5830

wduan@gwu.edu

Bin Gu

CBA 5.202, McCombs School of Business

The University of Texas at Austin, Austin, TX 78712

Phone: (512) 471-1582 Fax: (512) 471-0587

bin.gu@mccombs.utexas.edu

Andrew B. Whinston

CBA 5.202, McCombs School of Business

The University of Texas at Austin, Austin, TX 78712

Phone: (512) 471-7962 Fax: (512) 471-0587

abw@uts.cc.utexas.edu

March, 2008

Informational Cascades and Software Adoption on the Internet:

An Empirical Investigation

Acknowledgements

We are grateful to the Senior Editor Carol Saunders for helping us substantially improve the quality of our work. We also acknowledge the invaluable help of the Associate Editor and the three anonymous reviewers for their insightful comments and suggestions. The authors also thank Cynthia Beath, Hermant Bhargava, Vijay Mahajan, Jaejoon Han, the participants at the 2005 Marketing Science Conference, the 2005 Doctoral Internationalization Consortium in Information Systems, the 2005 International Conference on Information Systems (ICIS), the 2005 Workshop on Information Systems and Economics (WISE), and seminar participants at the University of Texas at Austin for valuable comments and suggestions. All errors are our own.

Author biographies

Wenjing Duan is an Assistant Professor in Information Systems and Technology Management at The George Washington University.

Bin Gu is an Assistant Professor in Information Systems at The University of Texas at Austin.

Andrew B. Whinston is the Hugh Roy Cullen Centennial Chair in Business Administration, Professor of Information Systems, Computer Science and Economics, Jon Newton Centennial IC2 Fellow, and Director of the Center for Research in E-Commerce at The University of Texas at Austin.

Informational Cascades and Software Adoption on the Internet:

An Empirical Investigation

ABSTRACT

Online users often need to make adoption decisions without accurate information about the product values. An informational cascade occurs when it is optimal for an online user, having observed others' actions, to follow the adoption decision of the preceding individual without regard to his own information. Informational cascades are often rational for individual decision making; however, it may lead to adoption of inferior products. With easy availability of information about other users' choices, the Internet offers an ideal environment for informational cascades. In this paper, we empirically examine informational cascades in the context of online software adoption. We find user behavior in adopting software products is consistent with the predictions of the informational cascades literature. Our results demonstrate that online users' choices of software products exhibit distinct jumps and drops with changes in download ranking, as predicted by informational cascades theory. Furthermore, we find that user reviews have no impact on user adoption of the most popular product, while having an increasingly positive impact on the adoption of lower ranking products. The phenomenon persists after controlling for alternative explanations such as network effects, word-of-mouth (WOM) effects, and product diffusion. Our results validate informational cascades as an important driver for decision making on the Internet. The finding also offers an explanation for the mixed results reported in prior studies with regard to the influence of online user reviews on product sales. We show that the mixed results could be due to the moderating effect of informational cascades.

Keywords: E-commerce, herding, informational cascades, decision making, network effects, word-of-mouth, software download, online communities, online user review.

Informational Cascade and Software Adoption on the Internet:
An Empirical Investigation

1. INTRODUCTION

As asserted by Eric Hoffer (1955): "When people are free to do as they please, they usually imitate each other." This leads to what is called *herd behavior*, i.e., *everyone is doing what everyone else is doing* (Banerjee 1992). Herding portrays various social and economic situations where an individual's or organization's decision making is markedly influenced by the decisions of others, such as in financial investment, technology adoption, firms' strategic decisions, political voting, and dining and fashion trends. For instance, when there are two restaurants next to each other, people often pick the one with more seats occupied; or despite mediocre reviews, a *New York Times* bestseller can sell well enough to continue as a bestseller (Bikhchandani et al. 1998). While *herding* describes a phenomenon in which individuals converge to a uniform social behavior (Bikhchandani et al. 1998), extant research has suggested a variety of mechanisms underlying such observations of convergent behavior among decision makers (Bikhchandani et al. 1992). Informational cascades, sanctions on deviants, positive payoff externalities, and conformity preference have been suggested as the primary mechanisms for herd behavior (Bikhchandani et al. 1992).

Herd behavior is particularly prominent in the IT industry. IT managers are known to follow each other in making IT investment decisions (Kauffman and Li 2003), and computer users often adopt popular software products, thus making them even more popular (Brynjolfsson and Kemerer 1996; Gandal 1994). The primary explanation provided in the extant literature for the observation of herd behavior in the IT industry is *network effects*, a form of positive payoff externalities.[2] Network effects refer to the idea that a product becomes more valuable as its user base expands (Katz and Shapiro 1994). For example, the value of an electronic payment system to a commercial bank increases with the number of banks that adopt the same system (Gowrisankaran and Stavins 2004). However, many researchers have argued that the significant network effects expected by academic researchers in the IT industry often fail to materialize (Liebowitz 2002). Even for products with network effects, there is often a limit as to the network size. For instance, it is well-known that peer-to-peer networks benefit from user participation, but too many users lead to network congestion and limit network effects on the margin (Asvanund et al. 2004). In such cases, a large network does not necessarily provide more value to potential adopters.

The concerns about the strength and existence of network effects in the IT industry require a fresh look at alternative causes of herd behavior. In this paper, we explore another major driver: *informational cascades* (Banerjee 1992; Bikhchandani et al. 1992; Li 2004). Informational cascades refer to the situation *"when it is optimal for an individual, having observed the actions of those ahead of him, to follow the behavior of the preceding individual without regard to his own information*" (Bikhchandani et al. 1992, pp. 994). Such a situation arises when decision makers have imperfect knowledge of the true value of a product so they infer its utility from observing actions of their predecessors. The influence of others' behavior

[2] Network effects and network externalities are often used interchangeably in the literature (Katz and Shapiro 1994). However, Liebowitz and Margolis (1994) pointed out that network effects are a much broader concept than network externalities. Network effects refer to the general phenomenon that "the benefit or surplus that an agent derives from a good [changes] when the number of other agents consuming the same kind of good changes" (Katz and Shapiro 1985). The term "network externalities," however, refers to a specific kind of network effects where "participants in the market fail to internalize these effects" (Liebowitz and Margolis 1994). That is, network externalities occur when market participants do not take the impact of their adoption decisions on others into consideration when they make the decisions. In this case, the market equilibrium is no longer socially optimal, and external intervention could be necessary. Liebowitz and Margolis (1994) noted that, in many cases, network effects do not lead to network externalities. This is because the owners of networks usually internalize the network effects by subsidizing initial adoptions. In networks without owners, coordination among network participants (e.g., industry standardization) can internalize the network effects.

could be so substantial that it dominates the influence of decision makers' own information. In this case, decision makers would imitate their predecessors without regard to their own information. Informational cascades are largely *informational* in nature (Berndt et al. 2003). As a result, despite the availability of close substitutes, informational cascades may lead to the dominance of one product or technology over another and sometimes may lead to the rejection of more efficient technologies (Abrahamson 1991). Thus, informational cascades offer an information-based explanation for the herd behavior (Bikhchandani et al. 1992; Bikhchandani and Sharma 2001; Li 2004), suggesting that individuals with noisy information may rationally ignore their own information and choose to follow the crowds.

Informational cascades could be particularly prominent on the Internet for two reasons. First, the vast amount of information available on the Web has created information overload among online users (Brynjolfsson and Smith 2000; Jones et al. 2004; Shapiro and Varian 1999). In particular, products ranging from software to electronics have become increasingly sophisticated, and assessment of their value and features often requires extensive knowledge about the products (Bakos 1991). Moreover, the number of competing products in each product category has grown exponentially as a result of the Internet's reach and the virtually unlimited shelf space of online retailers. Online shoppers often find that they lack the knowledge and time to make the optimal purchase decision out of dozens and sometimes hundreds of sophisticated competing products. In such cases, following others' choices could be the most efficient and rational way to make decisions, as suggested by the informational cascades theory. Second, the Internet and other digital channels provide much more information about other online users' choices and product popularity, therefore making informational cascades more feasible. Many online e-commerce websites display products according to their popularity, ranking items based on their previous sales performance. Amazon.com provides a *top sellers* list and posts each product's *sales ranking*. Software products listed at CNET Download.com are labeled as *popular* when the total number of downloads reaches top 50 in the previous week. Such popularity information is an indicator of the choices made by earlier adopters. The prevalence of popularity information on the Internet may drown out an online user's private value assessment and make her more prone to imitate earlier adopters' decisions. This allows informational cascades to start faster and to expand to a larger population (Bikhchandani et al. 1998). While informational cascades have often been observed in *local* environments (Bikhchandani et al. 1992), the pervasive use of the Internet and other information technologies extends their influence to the global stage.

The objective of this paper is to empirically examine the impact of informational cascades on product adoption on the Internet. A key empirical challenge in identifying the influence of informational cascades is to differentiate it from alternative factors that lead to herd behavior such as network effects, word-of-mouth (WOM) effects, and product diffusion. We address the challenge by noting two unique implications for informational cascades. First, alternative explanations predict that adoption increases with the number of previous adoptions, while informational cascades theory indicates that product adoption experiences distinctive leaps and drops with changes in popularity as measured by product adoption ranking. This distinction allows us to identify informational cascades when controlling for alternative explanations. Second, informational cascades suggest decision makers ignore information when following the crowds for popular products, but not when they adopt less popular products. Consistent with the predictions of the informational cascades literature, we find changes in product popularity indeed lead to distinctive leaps and drops in online users' software choices after controlling for network effects, WOM effects, and product diffusion. We also find user rating has no impact on user adoption of popular products, but its influence becomes significant for the adoption of less popular products. The influence consistently occurs in all eight software categories we studied. Our results indicate that informational cascades play a significant role in online users' decision making. The finding that user rating has an increasingly positive impact on less popular products also indicates that informational cascades moderate the influence of online user reviews. It provides a new explanation for the mixed results reported in prior studies, which often assume a uniform impact of online user reviews on product sales.

This paper adds to the IS research by offering a new perspective on the impact of information provided by the Internet on adoption decisions (Biros et al. 2002). The results have significant implications for designing customer-centric websites (Albert et al. 2004; Bapna et al. 2004) and devising business strategies for IT firms and online retailers. Traditionally during system design and analysis, more attention has been paid to user requirements and user acceptance, and less emphasis has been put on understanding how these design characteristics affect user behavior. This emphasis has been common because most IT systems are designed to meet specific technical requirements. However, as Hong et al. (2002) noted, a customer-centric website is not only an information system, but also a profit-generating channel. In this case, a deeper understanding of the relationship between design features, particularly product information features, and user behavior is important for the success of the websites (Palmer 2002). Our results indicate that design features providing product rankings have significant influence on online users' adoption decisions.

Our findings also suggest that for IT businesses to be successful, being superior in technology is not sufficient. As new technologies are constantly introduced to individual users and businesses, the market is overwhelmed with IT products, and users and businesses have neither the time nor the information to make well-informed adoption decisions. This provides a perfect environment for informational cascades, which may lead to the rejection of technically superior products. IT businesses need to take informational cascades into consideration when devising business strategies for their products. These strategies involve obtaining an early lead in product adoption and using product information strategically to influence informational cascades.

Our findings also suggest the abundance of online information does not necessarily benefit consumers. Informational cascades are known to generate suboptimal social allocation (Bikhchandani et al. 1998). By providing easily available product ranking information that enables informational cascades, the Internet may decrease rather increase consumer welfare.

The rest of the paper proceeds as follows. We discuss the related literature and research hypotheses in Sections 2 and 3, respectively. We describe data in Section 4, and develop and analyze the empirical model in Section 5. Section 6 discusses the results and implications, and we conclude the paper by addressing the limitations and identifying areas for future research.

2. LITERATURE REVIEW

The herding phenomenon has been widely observed in a variety of fields. Welch (1992) examined herd behavior in the market of Initial Public Offering (IPO) and found that investors mimic previous investors, regardless of their own substantive private information. Hirschleifer et al. (1994) find that herding influences not only investment choices, but also investors' decisions in acquiring information. Under some conditions, an investor will find it optimal to collect information about stocks that are followed by many investors. Bikhchandani et al. (2001) provided a comprehensive review of the recent theoretical and empirical research on herd behavior in financial markets. Herding is also observed beyond financial markets. Borenstein and Netz (1999) found that increased competition among airlines leads to even greater similarities in airline departure times. Kennedy (2002) indicated that, even though imitative introduction, on average, underperforms differentiated introduction, major television networks copy each other when introducing new programs. Simonsohn and Ariely (2004) tested their prediction of herd behavior in eBay's online auction. They found that bidders often engage in herd behavior in favoring auctions with more existing bids. Their results suggested that bidders interpret the number of existing bids as an informative signal of quality, even when the higher number of bids is simply caused by a lower starting price.

Herd behavior has also been observed in many cases of IT adoption in which investment decision makers follow the decisions of earlier adopters (Kauffman and Li 2003; Swanson and Ramiller 2004). Previous research has often attributed such imitative behavior to network effects (Au and Kauffman 2001; Gallaugher and Wang 2002; Kauffman et al. 2000). Network effects arise when adopters' payoff is increased by following the majority in adopting technology innovations (Katz and Shapiro 1985; Shapiro and Varian 1999). Network effects are expected to be prominent in many technology innovations. However, much imitative herd behavior in IT adoption and e-commerce cannot be single-handedly explained by network effects. Liebowitz (2002) noted that many e-commerce products do not even have the slightest trace of network effects. Li (2005) showed that bogus network effects can be stimulated through "cheap talk" and information asymmetry among different parties in a technology distribution channel.

One seminal contribution to the theoretical herding research is the development of informational cascades theory, which is a social learning mechanism introduced by Banerjee (1992) and Bikhchandani et al. (1992). In their framework, a sequence of individuals makes decisions with incomplete and asymmetric information. Individuals' information is private and inaccurate. Besides their own information, they also observe their predecessors' actions, although they do not observe their predecessors' decision-making process or relevant information used. Given their imperfect knowledge, sequential decision makers infer product values from the actions of their predecessors in combination with their private information. However, if consecutive predecessors choose a particular product, their influence on the following decision makers could be so strong that they imitate their predecessors regardless of their own private information. Bikhchandani et al. (1992) named the process *informational cascades*. Informational cascades are a special case of herd behavior. An informational cascade occurs when individuals ignore their private information when making a decision, whereas herd behavior takes place when all the individuals make an identical decision, although they do not necessarily ignore their private information (Smith and Sørensen 2000).

Informational cascades theory has often been recognized as the force behind observed herd behavior in an individual's or organization's decision making (Bikhchandani et al. 2001). Golder and Tellis (2004) showed that the impact of informational cascades is critical to the diffusion process in product life cycle. Berndt et al. (2003) demonstrated that the externalities arising from informational cascades significantly influenced the diffusion rate and market share of antiulcer drugs in the pharmaceutical market. Furthermore, Terlaak and King (2006) used the informational cascades theory as the foundation to test the bandwagon effect in adopting the ISO 9000 quality management standard.

A few recent IS studies considered the influence of informational cascades in IT adoption. Walden and Browne (2002) suggested that informational cascades play a significant role in influencing firms' adoption of electronic commerce technologies. Kauffman and Li (2003) proposed a theoretical framework for the herd behavior observed in IT adoption. They identified informational cascades theory as a new perspective to explain the dynamics of IT adoption. Li (2004), taking an explorative approach, focused on examining the influence of informational cascades in IT adoption and the business implications. He also explored various scenarios in IT adoption, in which informational cascades may interact with other mechanisms such as network effects and WOM effects. Konana and Balasubramanian (2005) referred to informational cascades theory in developing a Social-Economic-Psychological model to predict technology adoption of online banking. Informational cascades theory is also associated with the institutional theory of *mimetic isomorphism*, where institutions are observed to imitate one another in technology decision making (DiMaggio and Powell 1983; Tingling and Parent 2003; Swanson and Ramiller 2004). Both theories share the characteristics of peer influences and uncertainty in the decision-making process.

Complementing this line of IS research, we empirically explore informational cascades among online users and examine how online product information influences informational cascades. We differentiate our study on two fronts. First, we take the initial step to analyze herd behavior on the Internet, which, to our knowledge, has not been explored before. Second, the myriad of factors that lead to herd behavior have presented challenges to business practitioners. Without knowing the exact cause(s) of herd behavior, businesses have difficulties in exploring the opportunity or addressing the challenge presented by the mass herd behavior commonly observed. This is especially challenging for IT firms where informational cascades often intertwine with network effects, WOM effects, and other mechanisms in creating herd behavior. The objective of this paper is, therefore, to make a first step to address this challenge by identifying and measuring one key driver for the herd behavior – informational cascades - and differentiating it from other drivers.

3. RESEARCH HYPOTHESES

3.1. Research Context

Our empirical study is conducted in the context of software downloading at CNET Download.com (CNETD). CNETD presents an ideal environment for this study. All software programs listed on CNETD can be downloaded without charge; thus, the price effect on demand is controlled by default. In any software category, users are presented with a wide array of choices. In addition to the description of product features, two other types of product information are available on CNETD: information on previous users' choices and information on product evaluation. For information on previous users' choices, CNETD updates the cumulative number of downloads for each software program every day. It also reports the number of downloads for the most recent week. As to information on product evaluation, CNETD actively solicits user reviews for all programs and provides professional editor reviews on a small number (less than 10%) of the software programs. Visitors to CNETD can sort software on any of the attributes mentioned above. CNETD reports that the most often sorted attribute is number of downloads, which provides relative ranking of software downloads in a category. Figure 1 presents a screen shot of a typical product listing page on CNETD. The presence of both types of product information is not unique to CNETD. Many online e-commerce websites (e.g., Amazon.com and CircuitCity.com) offer similar ranking and rating information for individual products. By analyzing how the two types of information and their interactions affect online users' adoption decisions, we can identify informational cascades behavior.

Insert Figure 1 about here

3.2. Informational Cascades

Informational cascades theory considers a common phenomenon where a decision maker chooses among multiple competing products. The decision maker has two sources of information. One is her own information based on her knowledge of the products or her reading about the products. This information is often limited or imperfect, thus, she is uncertain of true product values. The other is the information derived from the adoption decisions of others. The decision maker combines the two sources of information to make the best decision. The influence of each source of information depends on their relative strength. If the decision maker's own information is very limited, she puts more weight on the information derived from observation of others' actions. If the decision maker is very knowledgeable about the products, the weight placed on others' decisions diminishes. The two sources of information may present conflicting signals from time to time and, when this happens, the decision maker follows the information source with more weight. An informational cascade arises when the influence of others' decisions outweighs the influence of the decision maker's own information, so that "*an individual's action does not depend on his private information signal*" (Bikchandani et al. 1992, pp. 1000). In this case, the decision maker simply follows her predecessors' decisions regardless of her private information. This in turn leads to a following of subsequent decision makers, making them adopt the same product with increasing momentum (Banerjee 1992; Bikhchandani et al. 1992). For example, in a simple setting in which customers make choices between two qualitatively equivalent products, it can be shown that a cascade will take place quickly when there have been two or more consecutive decisions of one "choice" over the other.[3]

[3] See Banerjee 1992, Bikhchandani et al. 1992, and Chamley 2004 for the proof of informational cascades models of two or more products.

The underlying message of informational cascades is that when people are making decisions based on their own information and their observation of others' actions, the influence of others' actions could be so substantial that it dominates the influence of their own information (Golder and Tellis 2004; Bikhchandani et al. 1992). Two key factors drive informational cascades: uncertainty in decision makers' own information and their ability to observe predecessors' decisions (Bikhchandani et al. 1992). Software downloading sites provide an excellent context for the study of informational cascades. Software is a typical form of experience goods where adopters face significant uncertainty in assessing values. Moreover, hundreds of software downloads are available for the same type of application, which further limits users' ability to assess their features and their values. To help online users make decisions, software downloading sites provide information on each software program's download counts and their ranking within each category. As predicted in informational cascades literature, such information represents recent adopters' choices and can significantly influence the followers' decision marking, especially when online users have limited information about product value and usability.

Informational cascades theory has a number of unique implications. First, it leads to a dramatic and distinctive change in the product adoption process. Before the start of informational cascades, adoption decisions are influenced by both private information and predecessors' decisions. The adoption process is gradual and noisy. But the process changes dramatically after the onset of informational cascades. Since individuals follow the behavior of their predecessors without regard to their own information, informational cascades lead to a sudden jump in the adoption process. Graphically, the start of informational cascades represents a tipping point where a gradual adoption process turns into frenzy (Bernt et al. 2003; Golder and Tellis 2003; Welch 1992).

The above discussion suggests that one of the distinctive characteristics of informational cascades is the presence of significant jumps in product adoption. In the extreme case, if none of the subsequent decision makers has better information, informational cascades will not stop once started and all of the later decision makers will adopt the same product (Bikhchandani et al. 1992). In reality, informational cascades do not last forever (Bikhchandani et al. 1998). The arrival of new information or more informed decision makers can stop informational cascades.

The jump in product adoption is triggered when the relative popularity, i.e., the sales ranking or download ranking in our context, of one product surpasses another. Consider a simple model where customers make sequential choices between two products, *A* and *B*.[4] Both products are of the same price, but one product is of higher quality than the other. Customers have no *ex ante* information about product quality, therefore both products are equally desirable to customers. Now assume that before each customer makes the purchase decision, she obtains a noisy signal with a probability $p>0.5$ that the signal reveals the true identity of the product with higher quality. The probability p, i.e., the strength of the signal, is the same across all customers. In this case, the first customer will choose the product based solely on her own signal. The second customer observes his own signal, but he can also infer the first customer's signal from her action. If both of the signals suggest the same product, the decision is straightforward. If the signals are in conflict, the second customer would be indifferent between the two products, since the strengths of the signals are the same. His choice will be determined by a random draw. The third customer's decision depends on the actions of the previous two customers. If both customers choose the same product, say *A*, the third customer infers that the first customer

[4] This example is adapted from Bikhchandani et al. (1998)

received a signal for *A* and the second customer is also likely to have received a signal for *A*. In this case, she will choose *A* regardless of her own signal, marking the beginning of informational cascades. To put it differently, *A*'s dominating market share provides a strong signal of *A*'s superiority over the other product, which triggers informational cascades. The aforementioned discussion shows that the observed *relative popularity* of the two products determines the timing and direction of informational cascades. When the popularity of one product outweighs the other, it sends a strong signal to later customers, and the signal dominates their own information.

It is important to note that popularity in the above discussion refers to the relative ranking of the two products rather than absolute sales. Once the popularity of one product exceeds the other and starts the informational cascades process, additional sales has no further impact on future product adoptions, because sales under informational cascades are not influenced by users' own information, and therefore have no informational value (Golder and Tellis 2004; Oh and Jeon 2007). This unique relationship between product ranking and product adoption distinguishes informational cascades from alternative explanations of herding phenomenon, such as network effects, WOM effects, or other mechanisms that link user valuation to product sales. These mechanisms often suggest that user adoption is a function of cumulative product sales. Since sales ranking is a variable derived from product sales, these mechanisms imply that sales ranking has no impact on product adoption after the influence of product sales is controlled for. Informational cascades theory, however, indicates the opposite. It suggests that user adoptions change with relative ranking of the products, but not with the absolute number of product sales, i.e., the number of software downloads in our research context. We therefore propose:

H1: Online users' choice of software products exhibits significant jumps and drops with changes in download ranking of the product.

Prior research on informational cascades theory considers individuals making adoption decisions between two products and notes that, once informational cascades start, individuals ignore their own private information to adopt the more popular product. The same principle applies in a multi-product environment (Banerjee 1992; Bikhchandani et al. 1992; and Chamley 2004. If individuals have the same preference, they will all choose to adopt the most popular product once informational cascades start. As a result, demand for the most popular product is expected to increase significantly while demand for all other products drop to zero. This winner-takes-all scenario, however, rarely materializes in reality because preferences among individuals differ. For example, some users prefer software with greater functionality; some prefer ease-of-use, while others have no particular preference. Such residual heterogeneity (Welch 2000) leads to variations in product adoption, which prevent the most popular product takes 100% of the market. To identify the implications of heterogeneity in consumer preference, we extend the previous example in *H1*. Consider a market with three products, *A*, *B* and *C*.[5] All the products are of the same price but different quality. Customers again have no *ex ante* information about product quality. Besides being differentiated on quality, the products differ in features that are observable to consumers. As customers have different preference for features, they may consider only a subset of the products. Some may consider two of the three products; and some may consider only one; and others may consider all three. Given the presence of three products, there are seven types of consumers as listed in Column 1 in Table 1. Now we analyze customer purchase decisions in an informational cascade. Without loss of generality, we assume that product *A* is the most popular product, product *B* the second most popular product and product *C* the least popular product. In an informational cascade, customers choose the most popular product within his or her consideration set. As a result, the most popular product (product *A*)

[5] This example is adapted from Bikhchandani et al. (1998)

sees the biggest increase in demand as it attracts all customers except those with preference that explicitly excludes product *A*. Column 2 of Table 1 shows that four out of the seven types of customers will buy product *A* in an informational cascade. The number two product (Product *B*) will see much less increase in adoption since it only attracts customers whose preference excludes the number one product (Product *A*). Table 1 shows that only two of the seven types of customers will choose to buy product *B*. Number three product (Product *C*) will see even less increase in adoption. In a market with three products, customers will purchase Product *C* if and only if their preference explicitly excludes Products *A* and *B*. This is clearly demonstrated in Table 1, where only one of the seven types of customers chooses Product *C*. The above discussion suggests that residual heterogeneity in customer preference moderates the winner-take-all situation of informational cascades, yet we will continue to see that high ranking products receive the most benefits under informational cascades. We therefore propose:

H2: Changes in download ranking lead to jumps and drops of higher magnitude for high ranking products.

3.3. Online User Reviews

Besides leading to jumps and drops in product adoption, another key implication of informational cascades is that decision makers ignore their own information when they make adoption decisions under the influence of information cascades (Bikhchandani et al. 1992). In our research context, a main information source for consumers is online user reviews submitted by fellow users.[6] These reviews provide a glimpse of others' evaluations of the product and help later adopters to form their own decisions.

[6] CNET also provides editor reviews on certain products. However, CNET editor reviews cover less than 10% of software programs. In addition, CNET tends to selectively review the most popular products, which limits CNET editor reviews' influence given that the product is already established. Moreover, CNET reviews do not change once they are posted for a software program. For most of the software in our study, the CNET reviews are either

The role of user reviews in influencing customer decision making has become an increasingly important topic for online businesses (Dellarocas 2003; Dellarocas et al. 2007), and we have seen a growing interest in understanding how these reviews influence users' adoption decisions. In spite of the widespread belief that the Internet may act as a huge "megaphone" in promoting product sales, literature has found mixed evidence on how online user reviews influence consumers' adoption decisions. Chevalier and Mayzlin (2006) showed a significant positive impact of online user reviews on relative online book sales for Amazon.com and BarnesandNoble.com. Chen et al. (2004) also used data from Amazon.com and yet attained different findings. Their results suggested that user ratings have no impact on sales. Pavlou and Dimoka (2006) took a more detailed look at the feedback comment text posted in eBay auction. They noted that most users only read very small portion of the feedback comments and the overall feedback score seems to be most influential. Their results indicate that the sentiment demonstrated in the comments has close relationship with the trust building, seller differentiation, and price premium.

Informational cascades theory offers a potential explanation of the mixed results discovered in earlier studies. The theory suggests that the relationship between online user reviews and product adoption is more complicated than previously considered. Instead of assuming that online user reviews have a uniform impact across products, informational cascades theory suggests that the impact is moderated by the level of informational cascades when users make adoption decisions. If a product is mainly adopted by users because of informational cascades, online user reviews have little impact on its sales. However, if a product is mainly adopted by users who apply their own information to make decisions, online user reviews can

---

posted before the study period or are never posted. In these cases, the impact of CNET reviews is captured by the time fixed effects as other time-constant variables in our empirical model. Therefore, we do not consider CNET review has major influence on product adoption.

significantly influence product sales. The nature of informational cascades suggests that products mainly adopted out of informational cascades are popular products, while users who choose to adopt less popular products are more likely to be driven by their own information. This suggests the following pair of hypotheses to test the presence of informational cascades.

H3: *Online user ratings have no impact on online users' choice of popular products.*

H4: *Online user ratings have an increasingly positive and significant impact on adoption of less popular products.*

4. DATA

4.1. Data Collection Methodology

Data for this study were collected from CNET Download.com (CNETD: http://www.download.com), which is part of the CNET networks. CNET networks provide reviews, news, and price information on technology products, as well as free software downloading and newsletters. CNETD is a library of over 30,000 free or free-to-try software programs for Windows, Macintosh, and handheld devices. Software programs are evaluated and categorized to facilitate user search. The list of available software programs can be sorted by total number or weekly number of downloads, software name, CNET rating, user rating, and the date added. Figure 2 shows the popularity of sort options used on CNETD listing pages, which has been published on the download.com webpage.[7] Download counts appear to be the most popular sort option (37%), suggesting that users either are concerned with network effects or place significant weight on their predecessors' choices. CNET's editorial staff reviews some of the software programs, with an emphasis on popular software programs. Reviews are summarized by ratings on a scale of *one* to *five*, with *one* being the worst and *five* the best.

[7] When we were collecting data, this figure was posted by CNETD on its website in the section of general introduction on posting software programs on CNETD.

CNETD also offers a widely used feedback system for online users to share their opinions and experiences. The user review system requests detailed comments, as well as an overall assessment indicated by a five-star user rating system.[8]

Insert Figure 2 about here

Our sample consists of eight software categories, which provides a diversified composition of the software programs. We ensure that our data cover a wide range of categories of software programs by including popular software categories belonging to different types of applications. There are a total number of 15 groups of software programs listed on CNETD. Each group has 8 to 22 categories. The categories we chose are the most popular software categories. The categories are adware & spyware removal, browsers, download managers, file compression, file sharing, Internet chat, media players and MP3 search tools. As shown later in Table 4, the eight categories represent close to 50% of the software in the most popular software list, indicating that they cover a large number of online user adoption decisions. We chose 8 categories to make sure they are large enough to provide a diversified coverage, while small enough to be manageable for data collection. Table 2 provides a summary of the eight software categories. The number of software programs listed in each category varies considerably from approximately 60 to 300. Such a variation reflects the idiosyncratic environment in a specific software category, which can be defined as a single *market*. We started collecting data in each category daily in November 2004. Every day we extracted the following information on every software program listed in each category: software name, description, date added, total download, last week download, CNET rating, number of user reviews, average user rating, and whether the

[8] CNETD originally used a binary (thumbs-up/thumbs-down) system for user reviews. At the end of January 2005, CNETD substituted the old systems with the five-star user-rating system, keeping everything else of interface the same. Since most of our data are collected in 2005, the analysis shown in this paper used the data using the current user review system.

software program has been labeled as *pop* (software is designated as *pop* if it climbs onto the most popular list) and *new* (software is defined as *new* for the first 15 days). We also collected software characteristics including operating system requirements, file size, publisher, license, and price if its license is *free-to-try*. Table 3 presents the variable definition, description, and explanation of measurement.

Insert Tables 2 and 3 about here

4.2. Key Variables

The key dependent variable in our empirical analysis is online users' choice of software products. Although we do not observe each individual's behavior, we can measure user choice collectively by using *weekly download market share* for each product in each individual *market* (i.e., in a single software category). Using market share instead of number of downloads has two main advantages. First, it removes unobservable influences, such as weekend and holiday effects, which affect the absolute number of downloads. Second, it standardizes the variance across software products, further facilitating the empirical analysis.

Our major independent variables include adoption decisions of recent predecessors and product information available online. The adoption decisions of recent predecessors are represented by download counts and download ranking, which highlights the popularity of the products in the market. CNETD prominently feature this information through the sorting functions of download counts. Product information provided by CNETD includes user reviews and professional editor reviews in addition to software features. CNETD actively solicits user reviews for all the products. CNETD also provides editor reviews for some software products (less than 10%). These reviews are usually posted a few months after the introduction of the software program. More details on variables and model specification will be discussed in section 5.

4.3. Descriptive Statistics

We start our analysis with descriptive statistics of the most popular software programs and show that they are not necessarily the best-rated products. This provides the first glimpse of the potential consequence of the informational cascades effect. CNETD provides rankings for the most popular titles in Windows each week, which includes the top 50 most-downloaded programs for the past week. We collected the most popular list each week from November 2004 to June 2005. For each piece of software on the list, we collected the following information: software name, description, this week's rank, last week's rank, weeks on the top 50 chart, this week's download, and total download. By aggregating data from November 2004 to June 2005, we find 106 unique entries of software programs that have appeared on the list. Table 4 shows the distribution of the software categories to which those 106 software programs belong.

Insert Table 4 about here

Table 5 presents the distributions of CNET ratings and user ratings for those 106 software programs. It shows that software with mediocre or even low CNET and user ratings can still make to the most-popular list. On a scale from one to five, 12 (13%) software programs that make to the list receive an average user rating below three. Similarly, 8 (11%) have CNET ratings below three. Our analysis of the most popular software programs suggests that CNET and user reviews may not be the main driving force behind online users' adoption decisions.

Insert Table 5 about here

Another approach to assess the potential presence of informational cascades is to consider how software adoption changes with download ranking on a weekly basis. Figure 3a plots the figure for software "Bitcomet" that is one of the most popular software in the category "File Sharing". The plot demonstrates significant jumps of market share with the increase of download ranking from #3 to #2, and to #1. In addition, the figure shows that market share exhibits higher jump when the download ranking changes from #2 to #1, compared with the increase from #3 to #2. This plot provide initial support for *H1* and *H2* that users' product choices exhibit significant jumps and drops with changes in download ranking, and that the changes have more impact on the sales of high ranking products. Figure 3b illustrates the relationship between download ranking and adoption for software "Star Downloader" that is listed in the category "Download Managers". It again shows significant jumps and drops in product adoption as the download ranking changes.

Insert Figure 3a and 3b about here

The initial findings described in this section seem to support the informational cascades theory that online users may adopt software programs according to previous users' choices. To test the hypotheses and fully answer our research questions, we develop an empirical model and conduct rigorous empirical analysis in the next section of the paper.

5. EMPIRICAL METHODOLOGY AND RESULTS

5.1. Empirical Model

Our major objective is to analyze the herd behavior in user adoption of software programs. We do not observe each user's download decision, but users' adoption decisions can be collectively measured by *download market share* in each individual market. We construct the measurement of market share for each individual software program, which reflects user adoption in a particular market. Let $i = 1....I$ index the software in a specific market. $DOWNLOAD_{it}$ is defined as the number of downloads of software i at time t. Hence, the download market share of software i at time t is:

$$S_{it} = \frac{DOWNLOAD_{it}}{\sum_{i=1}^{I} DOWNLOAD_{it}} \quad (1)$$

We employ the *multinomial logit* (MNL) market-share models as our basis of empirical analysis to explain the market shares (choices) of different products. MNL market-share models have been widely applied in marketing and economics literature to analyze the competitive structure and effects of advertisement and retail promotions (Cooper 1993). Since individual user choice underlies the process of market-share formation, the discussion of market-share models overlaps substantially with that of individual choice models (McFadden 1981). In addition, market-share models deal with market response over time as well as across competitors (Cooper 1993), which fits well into our situation. Following Nakanishi and Cooper (1974), the specification of the MNL market-share model is of the form:

$$\lg\left(\frac{S_{it}}{\tilde{S}_t}\right) = (\alpha_i - \overline{\alpha}) + \sum_{k=1}^{K} \beta_k \left(X_{ikt} - \overline{X}_{kt}\right) + \left(\varepsilon_{it} - \overline{\varepsilon}_t\right) \quad (2)$$

where S_{it} is the market share of the i-th product in a market of I products at time t, α_i is a parameter for the intrinsic value of product i, and X_{ikt} is the value of k-th exploratory variable, which may influence users' product choices, and ε_{it} is the error term. $\tilde{S}_t$ is the geometric mean of S_{it}, and $\overline{\alpha}$, $\overline{X}_{kt}$, and $\overline{\varepsilon}_t$ are the arithmetic means of α_i, X_{ikt}, and ε_{it}, respectively.

To apply the equation to our context, we need to first define markets. CNETD carries a large number of software programs that belong to many different categories, and these categories are mostly unrelated to each other. For instance, a user's choice of MP3 search software is independent from her choice of online chat software. Software programs within the same category, however, provide similar functionality and compete for the same group of users. We therefore treat each software category as an independent market and fit Equation (2) to each category separately. Technically, our estimation model is:

$$\lg\left(\frac{S_{it}}{\tilde{S}_{t,C}}\right)=\left(\alpha_i-\overline{\alpha}_C\right)+\sum_{k=1}^{K}\beta_{k,C}\left(X_{ikt}-\overline{X}_{kt,C}\right)+\left(\varepsilon_{it}-\overline{\varepsilon}_{t,C}\right) \tag{3}$$

where C denotes a particular software category, and the coefficients of the estimations are allowed to vary across different categories of software products. We note that we can rewrite the first term $\left(\alpha_i-\overline{\alpha}_C\right)$ in Equation (3) as α_i', which is just software-specific fixed effects. We also note that $\overline{\varepsilon}_{t,C}$ can be captured by time fixed effects $\alpha_{t,C}$, which vary across time periods for each software category, thus fully capturing the variation of any category time-varying variables. The above discussion suggests that we can rewrite Equation (3) as a linear estimation model:

$$\lg\left(\frac{S_{it}}{\tilde{S}_{t,C}}\right)=\alpha_i'+\alpha_{t,C}+\sum_{k=1}^{K}\beta_{k,C}\left(X_{ikt}-\overline{X}_{kt,C}\right)+\varepsilon_{it} \tag{4}$$

By stacking the software products data together in one category from different time periods, we construct a set of panel data to estimate Equation (4). Our panel data consist of weekly observations of software downloading and product information. Panel data can be used to obtain consistent estimators in the presence of unobserved variables (Wooldridge 2002). Equation (4) includes the software-specific fixed effects to capture the idiosyncratic and time-constant unobserved characteristics associated with each piece of software. The advantage of fixed effects estimation is that it controls for intrinsic software characteristics, which inherently affect market share. In addition, fixed effects estimation also allows the error term ε_{it} to be arbitrarily correlated with other explanatory variables, thus making the estimation results more robust. Equation (4) also includes the time-specific fixed effects that capture any influence on market share due to timing differences in each category.

To test *H1*, we include variable $WEEKLYRANK_{i,t-1}$. $WEEKLYRANK_{i,t-1}$ measures the most recent weekly download ranking of software *i* at time *t*. The coefficient represents the jump in adoptions a product experiences for each advance in product ranking. Note that a higher ranking corresponds to a lower value in $WEEKLYRANK_{i,t-1}$, the coefficient is expected to be negative, i.e., each retreat in product ranking leads to a drop in software adoption. To test whether the magnitude of jumps varies for products in different rank positions (*H2*), we add the square term of $WEEKLYRANK_{i,t-1}$ ($WEEKLYRANKSQ_{i,t-1}$). *H2* indicates that product adoption is a decreasing and convex function of product ranking. A ranking decrease for higher ranking products (e.g., a decrease from #1 to #2) leads a deeper drop in product adoption than for lower ranking products (e.g., a decrease from #4 to #5). The coefficient of the quadratic term is therefore expected to be positive. We note that a minor linear transformation of the rank variable and the square term could make the coefficients more intuitive (Please see Appendix 1 for a more detailed explanation of the transformation). Instead of including $WEEKLYRANK_{i,t-1}$ and $WEEKLYRANKSQ_{i,t-1}$, we consider ($WEEKLYRANK_{i,t-1}$-1) and its square term. For parsimony, we name the new variables $WEEKLYRANK\_1_{i,t-1}$ and $WEEKLYRANK\_1SQ_{i,t-1}$, respectively. The transformation makes the coefficient on $WEEKLYRANK\_1_{i,t-1}$ representing the impact of rank change on the most popular product and the coefficient on $WEEKLYRANK\_1SQ_{i,t-1}$ measuring how the impact varies for lower ranking products relative to the most popular product.

To test *H3* and *H4*, we introduce the $USERRATING_{i,t-1}$ and its interaction with product ranking $WEEKLYRANK_{i,t-1}$: *WEEKLYRANK*USERRATING*. Again, to present our results in the more intuitive way, we use $WEEKLYRANK\_1_{i,t-1}$ in place of $WEEKLYRANK_{i,t-1}$ for the interaction term (*WEEKLYRANK_1*USERRATING*).[11] The transformation again makes all the coefficients relative to the most popular product (rank #1). The coefficient on $USERRATING_{i,t-1}$ measures the influence of use reviews on user adoption of the *most popular* (rank #1) product, and the coefficient on the interaction term represents how the impact of online user reviews is moderated by informational cascades for lower ranking products relative to the most popular (rank #1) product. Informational cascades theory suggests that decision makers ignore information in adopting popular products. We, therefore, expect the coefficient on $USERRATING_{i,t-1}$ to be insignificant. Informational cascades are less likely to happen for less popular products, so the influence of user rating should increase for the lower ranking products. Thus, we expect the coefficient of the interaction term to be positive, which suggests that online user reviews have a significantly positive impact on less popular products. As a control, we include $USERRATINGD_{i,t-1}$, which specifies whether product *i* has received any user reviews. We also include $CNETRATINGD_{i,t-1}$ to indicate whether this product has received a CNET professional review.[9]

5.2. Informational Cascades, Network Effects, WOM, and Product Diffusion

To identify informational cascades, we need to control for other factors that influence online users' adoption decisions. As we mentioned earlier, many software products are expected to be subject to network effects (Gallaugher and Wang 2002; Kauffman et al. 2000), which also suggests that online users' adoption decisions are influenced by others' adoption decisions. The two effects need to be separated in order to identify the true influence of informational cascades. Moreover, prior studies on WOM and product diffusion also suggest alternative ways that product adoption decisions can be influenced by others' actions. WOM literatures indicate that exposure to WOM online and offline could significantly influence consumers' product adoption decisions (Godes and Mayzlin 2004). Product diffusion literatures (e.g., Bass 1969) find that consumer adoption is influenced by a product's current user base and the number of its potential users. As its user base expands, the adoption rate among potential users increases. In the

[9] We have tested alternative specifications by including the interaction term of CNET rating and download ranking to investigate whether the influence of CNET rating is moderated by informational cascades. We did not find any significant results, which indicate that in our research context, the editor's rating has no significant impact. Please also see footnote 6 for a detailed explanation on the nature of CNET rating and our modeling approach.

following sections, we discuss how to separate informational cascades from network effect, WOM effects, and product diffusion.

Control for Network Effects and WOM Effects: While both network effects and informational cascades indicate that adoption decisions are influenced by earlier adopters, there is a key difference between the two processes. Network effects indicate that the value of a product increases with the network size. In the context of software products, it means the value of a software increases with the total number of its users. Informational cascades, on the other hand, are information based. Informational cascades start when one product's popularity exceeds another and the information inferred from such popularity ranking overweighs individuals' own information. Different from network effects, increases in number of users have no impact on informational cascades if the increases do not change relative popularities between products. This is a crucial distinction between informational cascades and network effects.

Extant literature models network effects by assuming the adoption rate or the expected value of the software products increases with the size of a product's installed base (Brynjolfsson and Kemerer 1996). Previous research often uses the total number of the installed base to estimate the network effects (Gallaugher and Wang 2002; Brynjolfsson and Kemerer 1996). Following the same convention, we include variable $TOTALDOWNLOAD_{i,t-1}$, which measures the cumulative number of downloads of software i until time $t-1$, to control for the network effects on user software choices.

WOM communication is also often attributed as a leading cause for herd behavior (Bikhchandani et al. 1992; Li 2004). Extant innovation diffusion and information economics research suggests that WOM learning can significantly influence technology adoption and diffusion (Li 2004; Dellarocas et al. 2007). WOM is often associated with recent adoptions

(Elberse and Eliashberg 2003). Higher number of adoption indicates that more awareness of the product will be established through users who have adopted the product. In our context, we use the most recent weekly number of downloads (*LASTWEEKDOWNLOAD*$_{it}$) to control for the WOM effect.

Control for Product Diffusion: There are two diffusion processes for software downloads. First, a newly introduced software category, such as MP3 players, may gradually gain popularity among online users. To control for product diffusion at the category level, we use market share instead of the absolute number of downloads as the dependent variable. The market share model is not affected by the overall expansion or contraction of a product category, which eliminates the need to model diffusion at the category level. Second, within each category, individual software programs have their own diffusion processes. The product diffusion literature suggests that the level of adoption for each software product is influenced by a product's current user base and the number of its potential users. The larger the user base, the higher the adoption rate among potential users. This assumption is mathematically identical to the implication from the network effects theory. However, product diffusion theory also indicates that increases in user base carry a trade-off. With more and more users adopting the product, the number of potential users decreases. Consequently, the market share of new adoptions decreases as the product starts to saturate the market. This is an effect that we need to control for in assessing the true impact of informational cascades in the market. To control for this effect, we incorporate the Bass Diffusion Model into Equation (4).

The Bass Diffusion Model (Bass 1969) has been widely used in marketing (Mahajan et al. 2000) and economics (Berndt et al. 2003) for modeling new product diffusion. The basic Bass Diffusion Model indicates that product adoption at any given point is influenced by the current user base of a product and by potential users who have not adopted the product. The expression of Bass Diffusion Model of the number of new adopters at time t (Mahajan et al. 1993, p. 354) is:

$$n_i(t) = m_i \left[\frac{p_i (p_i + q_i)^2 e^{-(p_i + q_i)t}}{\left(p_i + q_i e^{-(p_i + q_i)t}\right)^2} \right] \tag{5}$$

In the above equation, m_i represents the total market potential for product i; and t denotes the age of the product since its first introduction; p_i stands for online users' inherent adoption rate to download the software; and q_i represents the increase in the adoption rate due to the increase in the user base.

Converting Equation (5) into the log market share is straightforward:

$$\lg \frac{n_i(t)}{n(t)} = \lg m_i - \lg p_i + 2\lg(p_i + q_i) - (p_i + q_i)t - 2\lg\left(1 + \frac{q_i}{p_i} e^{-(p_i + q_i)t}\right) - \lg n(t) \tag{6}$$

Then, applying the Taylor series to Equation (6),[10] we have:

$$\lg \frac{n_i(t)}{n(t)} = \lg m_i - \lg p_i + 2\lg(p_i + q_i) - (p_i + q_i)t - \frac{2q_i}{p_i} e^{-(p_i + q_i)t} - \lg n(t) \tag{7}$$

The first three items of the above equation are unique to each product and will be captured by product fixed effects. The last item represents the total number of downloads for the product category at time t, which will be captured by time fixed effects. The two items in the middle suggest that the log market share of a product depends on the age of the product and that the relationship contains both linear and nonlinear components. To facilitate linear regression, we use the conventional quadratic term to approximate the nonlinear component. That means we can add product age (AGE_{it}) and the quadratic term of product age ($AGESQ_{it}$) into the empirical model to control for product diffusion. Prior research shows that individual products often diffuse at different speeds, with some reaching market saturation faster than others (Mahajan et al. 2000). To model the differences in diffusion processes, ideally we need to allow coefficients

---

[10] The Taylor series approximation is valid when $p_i e^{(p_i + q_i)t} \gg q_i$, which can be readily satisfied when t increases. We estimate the model for samples encompassing different time periods and have obtained qualitatively similar results.

on both AGE_{it} and $AGESQ_{it}$ to vary across products. However, this approach does not provide enough degrees of freedom for our empirical analysis. Thus, we take a compromising approach that allows the coefficient on AGE_{it} to vary across products but keeps the coefficient on $AGESQ_{it}$ fixed for products in the same category. Our approach essentially assumes that each software program may diffuse at a different speed, but the shape of the diffusion process is the same across all software in the same category. We consider this to be a reasonable compromise, as it captures the first-order differences in diffusion processes while maintaining enough degrees of freedom for our analysis.

In addition to network effects, WOM effects, and product diffusion, other factors may be correlated with online user adoption of a software program. We control for these factors by including a dummy variable $POPD_{it}$ that indicates whether software *i* is labeled as *pop* at time *t*; $NEWD_{it}$ that specifies whether software *i* is labeled as *new* at time *t,* since users may be more interested in new software programs; and $NUMSOFTWARE_{it}$ that is the total number of software programs listed in this category at time *t* to control for the competition in a particular software category *t.*

5.3. Panel Data Fixed Effects Estimation Results

We estimate our empirical model using the panel data of weekly observations. We use weekly data because it covers longer time periods and many of our key variables are updated weekly.[11] We construct the unbalanced panel data for the eight software categories during the time period between February 2005 and June 2005.[12] Tables 6a and 6b show the descriptive statistics of the key variables. Notice that the weekly market share can be as high as 75% and as

[11] We also test our model using daily observations at different time periods, which generates qualitatively similar results; the report is available upon request.
[12] CNETD changed its user review system in late January 2005. To make sure our results are not affected by this change, we restrict our data analysis to data collected after that change.

low as almost zero, indicating market dominance by a small number of software programs. The high discrepancy in weekly market share illustrates users' preferences for popular software programs, suggesting the existence of herding. Tables 7a and 7b present the correlation matrix for the pooled weekly data of the eight software categories.

Insert Tables 6a, 6b, 7a and 7b about here

The fixed effects panel data estimation results are shown in Table 8 with centered $lgWEEKLYSHARE_{it}$ as the dependent variable. The coefficient on $WEEKLYRANK\_1_{i,t-1}$ is significantly negative across all eight categories. As we mentioned earlier, this coefficient measures the marginal impact of rank changes on the most popular product. The negative coefficient supports hypothesis *H1*, indicating that user adoptions see a significant jump when a product's relative ranking changes from the second place to first in popularity. *H2* is also strongly supported in all but one (File Compression) product category with a significantly positive coefficient of $WEEKRANK\_1SQ_{i,t-1}$, indicating that higher ranking products will see a greater increase in adoption as a result of increase in download ranking. The coefficient on $USERRATING_{i,t-1}$ has no consistent significant impact across the eight categories. As we mentioned earlier, the coefficient on $USERRATING_{i,t-1}$ measures the influence of user rating on the adoption of the most popular product (rank #1). The insignificance of the coefficient supports informational cascades theory that users ignore product information when they follow others to adopt the most popular product (*H3*). We find significantly positive coefficients of *WEEKLYRANK_1*USERRATING* in all eight categories. This suggests that the influence of online user reviews is moderated by informational cascades. User reviews have no impact for popular products, but have significant influence on users' choice for less popular products, supporting *H4*. The results also show that $TOTALDOWNLOAD_{i,t-1}$ is significantly positive for

Internet Chat, but has no impact for other categories, which is consistent with our expectation that software programs in *Internet Chat* have strong network effects. We also note the significant positive coefficient of $LASTWEEKDOWNLOAD_{it}$, suggesting the potential influence of WOM and informational value contained in recent week's downloads. Furthermore, we note that the variable $POPD_{it}$ has significant positive relationships with the market share for all eight categories, which complement our test of informational cascades using weekly rank.

Insert Table 8 about here

Recent econometrics development suggests that panel data often involve time series features that have not been considered in traditional estimation techniques (Wooldridge 2002). Particularly in our study, the weekly data may be autocorrelated. To ensure the robustness of our results, we incorporate the time series feature into our panel data model. Assuming a time series process with autocorrelation residuals, we estimate our panel data using a fixed-effects technique with autocorrelated (AR) disturbances. We also test alternative specifications such as estimating the model weighted by the number of user reviews, and estimating the model under heteroskedasticity robust errors. In all above discussed cases, the fixed-effects estimation gives qualitatively similar results. We thus omit the detailed statistics report here.

6. DISCUSSION, LIMITATIONS, AND FUTURE RESEARCH

In this study, we provide a systematic analysis of informational cascades by empirically investigating two sets of unique characteristics of informational cascades, namely, the impact of changes in ranking on product adoption and the lack of impact of product reviews on the adoption of popular products. We conduct our analysis in the context of online software downloading that illustrates the overwhelming amount of choices available in today's online market and the extent to which online users have access to others' adoption decisions and product review information. While the software downloading market potentially represents an extreme case of information overload and limited product knowledge, its key characteristics are shared by many other e-commerce sites. For example, Amazon.com lists 3,285 cameras in the category of “Point-and-Shoot Digital Cameras,” a choice set that is daunting to even the most sophisticated shoppers. To assist online users’ purchase decisions, Amazon.com provides a variety of sales ranking data, product information, and user reviews. Moreover, Amazon.com features top seller lists prominently on the category home page and updates the list hourly as a main vehicle to attract shoppers’ attention. Similar approaches have been adopted by other major online retailers. A quick survey shows that online stores ranging from Buy.com to iTunes.com provide others’ adoption decisions to facilitate users’ purchase decisions, and most of the online retailers prominently feature the best-selling products on their homepages. Our analysis of the software downloading market provides us with a detailed understanding of how information about others’ adoption decisions could influence subsequent adopters’ decisions in the online markets and how it moderates the influence of online product information and user reviews.

By analyzing our panel data using a fixed-effects market share model, our results reveal that online users’ choice of software is heavily driven by change in download ranking and popularity information after controlling for download counts and product review information. Our findings are consistent with the predictions of the informational cascades literature that individuals are remarkably influenced by the information inferred from others’ behavior.

Our results also shed light on the intriguing relationship between product information and informational cascades. We find that online user ratings have no impact on online users’ choice of popular products, whereas ratings have a significant positive impact on the adoptions of less popular products. This is again consistent with the informational cascades literature, which suggests that consumers ignore product information during informational cascades in adopting popular products. Our result also suggests a complicated relationship between online user reviews and product adoption decisions, which has not been addressed in previous literature and calls for extensions in future research.

In addition, our analysis separates informational cascades from network effects that often affect software adoption decisions (Brynjolfsson and Kemerer 1996). Both network effects and informational cascades lead to herd behavior, although the underlying causes of herding are different. Under network effects, user adoptions increase with increases in the overall installed user base of the software. Under informational cascades, user adoptions are driven by changes in relative ranking between products. We separate these two effects in this study by including user adoption to capture the network effects and relative ranking to capture the informational cascades effect. We find that network effects exist for certain products that can generate direct external benefits, while informational cascades have significant and consistent influence across all products. Previous IS research often attributes observed herd behavior to network effects without considering informational cascades, and thus might overstate the significance of network effects. Our analysis indicates that, while informational cascading is less known, it exerts a ubiquitous and dominant impact on a user's adoption decision.

While our analysis focuses primarily on user adoption of consumer-oriented products, it has implications for corporate IT adoption as well. Prior studies have indicated that informational cascades theory could be applied to the corporate IT adoption environment (Kauffman and Li 2003; Li 2004; Terlaak and King 2006). These studies consistently found that IT managers are more likely to adopt popular information systems, and one of the drivers of the phenomenon is informational cascades. Given the high stakes involved and the earlier findings that IT managers are influenced by informational cascades, our results suggest that a corporation can mitigate the situation by providing more information to IT decision makers. This could involve providing better training, providing access to professional product research reports, and/or hiring better IT managers who have extensive product and industry knowledge.

This paper has a number of limitations. To facilitate discussion, we provide in Appendix 2 a list of assumptions used in the paper, their impact on the paper, and possible future research directions to address these assumptions. First, in this study, we only consider the product information users can observe on CNETD, without addressing other available information resources, such as other third-party review sites, software magazines, and word-of-mouth from family and friends. While CNETD is the dominant software downloading site and often the only online information source for software products, it is known that WOM from trusted friends and family members can be an important source of information for online users. Incomplete information about users' information sources reduces the accuracy of analysis. Future research is necessary to examine the influence of multiple information resources and the limitation of information provided by a firm's own site. In addition, we implicitly assume that online users pay attention to sales ranking information, online product information, and user reviews in our model. We do not know, however, whether they really do or how differences in websites influence their interests and use of such information. A survey of online users with regard to their information may allow future research to generate a better understanding of the impact of online product information.

Second, the use of free or free-to-try products may affect the generalizability of our results. Online users may take decisions about free or free-to-try software programs less seriously compared to decisions about more expensive technical innovations. We investigate this issue and find that free or free-to-try software programs often require significant commitments from the users. For example, MP3 software is often used to manage music files, and the users need to input extensive information about these music files and put in significant effort to process these files. The time and effort spent will be wasted if the user later decides to remove the program. Similarly, Chat software often requires a user to set up a new account, import the contact list, and inform her friends of the transition to the new program. If the user decides to abandon the program later, she would expect another messy transition. In addition, many programs require users to learn new interfaces and explore new functionality. All those efforts required of software users represent a significant cost for using these programs. From this perspective, free-to-download software programs are not substantially different from other products that online users have to purchase.

Third, our empirical analysis cannot fully distinguish rational and irrational herding. Herding that arises from informational cascades is *rational* in the sense that decision-makers incorporate their predecessors' actions into their own decision-making process. Informational cascades occur when the decisions of others are strong enough to outweigh individual decision-makers' private information. However, the same type of herd behavior can also be generated from totally *irrational* behavior when decision makers just ignore the information available to them. Interestingly, our results provide some hints that the observed herd behavior is likely to be to the result of the rational herding. Our results suggest that consumers pay more attention in adoption of less popular products, suggesting certain rationality in the decision making process Separating rational and irrational herd behavior is certainly an interesting extension to this study.

Fourth, this study characterizes how informational cascades moderate the influence of product information on online user choices. In the process, we assume product information and user reviews as given. In reality, product information from user reviews is generated by users themselves. This information generation process is dynamic as online users influence each other in forming opinions about a product. Recent research has shown that online users could be subject to both persuasion biases and extremism in that they are more likely to post reviews if they feel strongly (positively or negatively) about a product (Dellarocas and Narayan 2006; Hu et al. 2006). Li and Hitt (2008) found that another mechanism, self-selection, may produce biased user ratings as well. However, we do not know the causes of the divergence and how they affect the informational cascades process. Future research targeting the dynamics of user reviews and their interaction with informational cascades would be fruitful.

Fifth, we consider network effects and informational cascades as two distinct mechanisms that lead to herd behavior. However, informational cascades and positive network effects may complement each other in various situations. For a product with positive network effects, the herding observed due to informational cascades will be further accelerated by network effects, resulting in more followers adopting the product. Therefore, for products with positive network effects, it is even more important to take up the market in the early stage in order to fully leverage the influence of informational cascades. An extension to the current research would be valuable to examine the complementarities between informational cascades and network effects in influencing the product adoption and diffusion process.

Finally, we make a few assumptions in deriving the empirical model. For example, in our effort to control for product diffusion, we assume the shape of the diffusion curve to be the same across different products, although we allow diffusion speed to vary. We also choose a particular way to incorporate the Bass Diffusion Model into the regression model to maintain its linear structure. The choice of a linear estimation model facilitates estimation of key coefficients. However, it sacrifices some of the richness of the diffusion model. Future research may benefit from using nonlinear models for a better control of product diffusion.

REFERENCES

Abrahamson, E. "Managerial Fads and Fashions: the Diffusion and Rejections of Innovations," *Academy of Management Review* (16:3), 1991, pp. 586-612.

Albert, T. C., Goes, P. B., and Gupta, A. "GIST: A Model for Design and Management of Content and Interactivity of Customer-Centric Web Sites," *MIS Quarterly* (28:2), 2004, pp. 161-181.

Asvanund, A., Clay, K., Krishnan, R., and Smith, M. "An empirical analysis of network externalities in peer-to-peer music sharing networks," *Information Systems Research* (15:2), 2004, pp. 155-174.

Au, Y. A., and Kauffman, R. J. "Should We Wait? Network Externalities, Compatibility, and Electronic Billing Adoption," *Journal of Management Information Systems* (18:2), 2001, pp. 47-63.

Bakos, J. Y. "A Strategic Analysis of Electronic Marketplaces," *MIS Quarterly* (15:3), 1991, pp. 295-310.

Banerjee, A. V. "A Simple Model of Herd Behavior," *Quarterly Journal of Economics* (107:3), 1992, pp. 797-817.

Bapna, R., Goes, P., Gupta, A. and Jin, Y. "User Heterogeneity and its Impact on Electronic Auction Market Design: An Empirical Exploration," *MIS Quarterly* (28:1), 2004, pp. 21-43.

Bass, F. M. "A New Product Growth Model for Consumer Durables," *Management Science* (15), 1969, pp. 215-227.

Berndt, E. R., Pindyck, R. S., and Azoulay, R. "Consumer Externalities and Diffusion in Pharmaceutical Markets: Antiulcer Drugs," *Journal of Industrial Economics* (LI:2), 2003, pp. 243-269.

Bikhchandani, S., Hirshleifer, D., and Welch, I. "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy* (100:5), 1992, pp. 992-1026.

Bikhchandani, S., Hirshleifer, D., and Welch, I. "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascasdes," *Journal of Economic Perspectives* (12:3), 1998, pp. 151-170.

Bikhchandani, S., and Sharma, S. "Herd Behavior in Financial Markets," *IMF Staff Papers* (47:3), 2001, pp. 279-310.

Biros, D. P., George J. F., and Zmud, R. W. "Inducing Sensitivity to Deception in Order to Improve Decision Making Performance: A Field Study," *MIS Quarterly* (26:2), 2002, pp. 119-144.

Borenstein, S., and Netz, J. "Why Do All The Flights Leave at 8 am?: Competition and Departure-Time Differentiation in Airline Markets," *International Journal of Industrial Organization*, 1999, pp. 611-640.

Brynjolfsson, E., and Kemerer, C. F. "Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market," *Management Science* (42:12), 1996, pp. 1627-1647.

Brynjolfsson, E., and Smith, M. D. "Frictionless Commerce? A Comparison of Internet and Conventional Retailers," *Management Science* (46:4), 2000, pp. 563-585.

Chamley, C. P. *Rational Herds: Economic Models of Social Learning*, Cambridge University Press, Cambridge CB2 2RU, UK, 2004.

Chen, P.-y., Wu, S.-y., and Yoon, J. "The Impact of Online Recommendations and Consumer Feedback on Sales," in *Proceedings of the 25th International Conference on Information Systems,* Washinton, D.C., 2004, pp. 711-724.

Chevalier, J. A., and Mayzlin, D. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research* (XLIII), August 2006, pp. 345-354.

Cooper L. G. " Market-Share Models," J. Eliashberg and G.L. Lilien, eds. *Handbooks in Operations Research and Marketing Science* (vol. 5), Elsevier Science Publishers B. V., 1993, Amsterdam: North-Holland, pp. 259-314.

Dellarocas, C. N. "The Digitization of Word-of-Mouth: Promise and Challenges of Online Feedback Mechanisms," *Management Science* (49:10), 2003, pp. 1407-1424.

Dellarocas, C. N., Awad, N. F., and Zhang, X. "Exploring the value of Online Product Reviews in Forecasting Sales: The Case of Motion Pictures," *Journal of Interactive Marketing* (21:4), 2007, pp. 23-45.

Dellarocas, C. N., and Narayan, R. "A Statistical Measure of a Population's Propensity to Engage in Post-Purchase Online Word-of-Mouth," *Statistical Science* (21:2), 2006, pp. 277-285.

Dimaggio, P. J., and Powell, W. W. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields," *American Sociological Review* (48), 1983, pp. 147-160.

Elberse, A. and Eliashberg, J. "Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures." *Marketing Science* (22:3), 2003, pp. 329-354.

Gallaugher, J. M., and Wang, Y.-M. "Understanding Network Effects in Software Markets: Evidence from Web Server Pricing," *MIS Quarterly* (26:4), 2002, pp. 303-327.

Gandal, N. "Hedonic Price Indexes for Spreadsheets and an Empirical Test of the Network Externalities Hypothesis," *RAND Journal of Economics* (25:1), 1994, pp. 160-170.

Godes, D., and Mayzlin, D. "Using Online Conversations to Study Word of Mouth Communication," *Marketing Science* (23:4), 2004, pp. 545-560.

Golder, P. N., and Tellis, G. J. "Growing, Growing, Gone: Cascades, Diffusion, and Turning Points in the Product Life Cycle," *Marketing Science* (23:2), 2004, pp. 207-218.

Gowrisankaran, G., and Stavins, J. "Network Externalities and Technology Adoption: Lessons from Electronic Payments," *RAND Journal of Economics* (35), 2004, pp. 260–276.

Hirshleifer, D., Subrahmanyam, A., and Titman, S. "Security Analysis and Trading Patterns When Some Investors Receive Information before Others," *Journal of Finance*, 49(5), 1994, pp. 1665-1698.

Hoffer, E. *The Passionate State of Mind*, Harper, New York, 1955.

Hong, W., Tam, K., and Yim, C. "E-Service Environment: Impacts of Web Interface Characteristics on Consumers' On-line Shopping Behavior," in *E-Service: New Directions in Theory and Practice*, R. Rust and P. Kannan (Eds), M. E. Sharpe, Armonk, NY, 2002, pp.108-130.

Hu, N., Pavlou, P. A., and Zhang, J. "Can Online Reviews Reveal a Product's True Quality? Empirical Findings and Analytical Modeling of Online Word-of-Mouth Communication," in *Proceedings of the 7th ACM conference on Electronic commerce,* Ann Arbor, Michigan, 2006, pp. 324-330.

Jones, Q., Ravid, G., and Rafaeli, S. "Information Overload and the Message Dynamics of Online Interaction Spaces: A Theoretical Model and Empirical Exploration," *Information Systems Research* (15:2), 2004, pp. 194–210.

Konana, P., and Balasubramanian, S. "The Social–Economic–Psychological Model of Technology Adoption and Usage: An Application to Online Investing," *Decision Support Systems* (39:3), 2005, pp. 505-524

Katz, M. L., and Shapiro, C. "Network Externalities, Competition, and Compatibility," *American Economic Review* (75:3), 1985, pp. 424-440.

Kauffman, J. R., and Li, X. "Payoff Externalities, Informational Cascades and Managerial Incentives: A Theoretical Framework for IT Adoption Herding," Working Paper WP 03-18, Management Information Systems Research Center, University of Minnesota, Minneapolis, 2003.

Kauffman, R., McAndrews, J., and Wang, Y.-M. "Opening the 'Black Box' of Network Externalities in Network Adoption," *Information Systems Research* (11:1), 2000, pp. 61-82.

Katz, M. L., and Shapiro, C. "Systems Competition and Network Effects," *The Journal of Economic Perspectives* (8:2), 1994, pp. 93-115.

Kennedy, E. R. "Strategic Fads and Competitive Convergence: An Empirical Test for Herd Behavior in Prime-Time Television Programming," *The Journal of Industrial Economics* (L:1), 2002, pp. 57-84.

Li, X. "Informational Cascades in IT Adoption," *Communications of the ACM* (47:4), 2004, pp. 93-97.

Li, X. "Cheap Talk and Bogus Network Externalities in the Emerging Technology Market," *Marketing Science* (24:4), 2005, pp. 531-543.

Li, X., and Hitt, L. "Self Selection and Information Role of Online Product Reviews." forthcoming in *Information Sysstems Research*

Liebowitz, S. J. *Rethinking the Network Economy: The True Forces That Drive the Digital Economy*, Amacom Publishing: Saranac Lake, NY, 2002.

Liebowitz, S. J., and Margolis, S. E., "Network Externality: An Uncommon Tragedy," *The Journal of Economic Perspectives* (8:2), 1994, pp. 133-150.

Mahajan, V., Muller, E. and Bass, F. M. "New-Product Diffusion Models," in *Handbooks in Operations Research and Management Science: Marketing*, edited by J. Eliashberg and G. L. Lilien. Elsevier Science Publishers B. V. 1993.

Mahajan, V., Muller, E. and Wind, Y. *New-Product Diffusion Models*, Kluwer Academic Publishers, Boston, MA, 2000.

McFadden, D. “Econometric Models of Probabilistic Choice,” Charles Manski, Daniel McFadden, eds. *Strutural Analysis of Discrete Data with Econometric Applications.* MIT Press, 1981, Cambridge, MA, pp. 198-272.

Nakanishi, M., and Cooper, L. G. “Parameter Estimation for a Multiplicative Competitive Interaction Model-Least Squares Approach,” *Journal of Marketing Research* (11), 1974, pp. 303-311.

Oh, W. and Jeon, S. “Memebership Herding and Network Stability in the Open Source Community: The Ising Perspective,” *Management Science* (53:7), 2007, pp. 1086-1101.

Palmer, J. W. “Web Site Usability, Design, and Performance Metrics,” *Information Systems Research* (13:2), 2002, pp. 151-167.

Pavlou, P. A. and Dimoka, A. “The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation,” *Information Systems Research* (17:4), 2006, pp. 391-412.

Shapiro, C., and Varian, H. *Information Rules*, Harvard Business School Press, Boston, 1999.

Simonsohn, U., and Ariely, D. “e-Bay's Happy Hour: Non-rational Herding in on-line Auctions,” Working Paper, The Wharton School, University of Pennsylvania, 2004.

Smith, L., and Sorensen, P. N. “Pathological Outcomes of Observational Learning,” *Econometrica* (68:2), 2000, pp. 371-398.

Swanson, B. E., and Ramiller, N. C. “Innovating Mindfully with Information Technology,” *MIS Quarterly* (28:4), 2004, pp. 553-583.

Terlaak, A., and King A. A. “Choosing Whom to Follow: Profitability Expectations and the Role of Smaller Organizations in Information-Based Adoption Processes,” Working Paper, Tuck School of Business at Dartmouth, 2006.

Tingling P., and Parent M. “Mimetic Isomorphism and Technology Evaluation: Does Imitation Transcend Judgment?” *Journal of Association for Information Systems* (3), 2002, pp. 113-143.

Walden, A. E., and Browne, J. G. “Information Cascades in the Adoption of New Technology,” in *Proceedings of the 23th International Conference on Information Systems*, Barcelona, Spain, 2002, pp. 435-443.

Welch, I. “Sequential Sales, Learning, and Cascades,” *The Journal of Finance* (47:2), 1992, pp. 695-732.

Welch, I. "Herding Among Security Analysts," *Journal of Financial Economics* (58:3), 2000, pp. 369-396.

Wooldridge, J. M. *Econometric Analysis of Cross Section and Panel Data*, the MIT Press, Cambridge, MA, 2002.

Appendix 1: Transformation of the Weekly Rank Variable and the Interpretation

To test *H1* and *H2*, we include the most recent weekly download ranking of software *i* at time *t* $WEEKLYRANK_{i,t-1}$ and its quadratic term $WEEKLYRANKSQ_{i,t-1}$. Instead of including $WEEKLYRANK_{i,t-1}$ and $WEEKLYRANKSQ_{i,t-1}$, we consider ($WEEKLYRANK_{i,t-1}$-1) and its square term. That is, we transformed the original rank (*1, 2, 3*…) to (*0, 1, 2*…). The simple linear transformation does not change the regression model, yet makes the interpretation of the coefficients easier and more intuitive. Given the transformation, the impact of weekly ranking on product sales can be expressed as follows: $\beta_1(WEEKLYRANK_{i,t-1} - 1) + \beta_2(WEEKLYRANK_{i,t-1} - 1)^2$. The first derivative of the expression shows that the marginal influence of change in weekly rank is $\beta_1 + 2\beta_2(WEEKLYRANK_{i,t-1} - 1)$. This gives an intuitive explanation of coefficient β_1, which stands for the marginal influence of rank changes for the most popular product (rank #1), when the second term of the equation $(WEEKLYRANK_{i,t-1} - 1) = 0$.

To test *H3* and *H4*, we introduce the $USERRATING_{i,t-1}$ and its interaction with product ranking $WEEKLYRANK_{i,t-1}$: *WEEKLYRANK*USERRATING*. Similar as above, to present our results in the most understandable way, we use $WEEKLYRANK\_1_{i,t-1}$, which equals to ($WEEKLYRANK_{i,t-1}$ -1) in place of $WEEKLYRANK_{i,t-1}$ for the interaction term (*WEEKLYRANK_1*USERRATING*) . That is, the impact of user rating on product adoption can be expressed as $\beta_3 USERRATING_{i,t-1} + \beta_4(WEEKLYRANK_{i,t-1} - 1)USERRATING_{i,t-1}$. The expression allows the influence of user ratings to vary with product ranking. The first derivative of the expression relative to $USERRATING_{i,t-1}$ shows that the marginal influence of user rating is $\beta_3 + \beta_4(WEEKLYRANK_{i,t-1} - 1)$. For the most popular product (rank #1), the influence of user rating is simply β_3, given $\beta_4(WEEKLYRANK_{i,t-1} - 1)USERRATING_{i,t-1} = 0$. That is, the transformation allows the coefficient on $USERRATING_{i,t-1}$ to measure the influence of use reviews on adoption of the *most popular* (rank #1) product, and the coefficient on the interaction term represents how the impact of online user reviews is moderated by informational cascades for lower ranking products relative to the most popular product.

Appendix 2: Summary of Key Assumptions, Limitations, and Future Research

Assumption	Limitation	Future Research
1. Online users pay attention to the information presented on CNETD, and adoption decisions are mainly driven by information on CNETD.	We do not observe online users' information collection and usage process. In addition, online users may also get information from their friends and other popular press sources.	To survey online users with regard to their information usage and examine the influence of multiple resources
2. User adoption of free-to-try software is similar to their purchase behavior of other online products.	Users need to incur significant learning costs and switching costs to adopt a software product. From this perspective, learning costs are the price users pay for the software. However, users may respond differently to prices versus costs of adoption.	To consider herd behavior in other contexts where price is a factor for consideration.
3. Herd behavior tested is caused by a *rational* informational cascading process.	*Rational* and *irrational* herding can both occur, which may lead to similar outcomes.	To differentiate *rational* and *irrational* herding
4. Product review information is taken as exogenous.	Product reviews, especially online user reviews, are dynamic.	To explicitly incorporate the dynamics of online user reviews and its influence into the product diffusion process
5. Informational cascades and network effects are two separate mechanisms that lead to herd behavior.	There may exist significant complementarities between informational cascades and network effects.	To model complementarities between informational cascades and network effects.
6. The shape of the diffusion process is the same across all software products. Bass Diffusion Model is incorporated into the empirical model in a log linear form.	Details of Bass diffusion model, such as differences in diffusion shapes, are lost.	To relax the assumption on parameter and distribution specifications of the product diffusion process, potentially with a nonlinear model

TABLES AND FIGURES

Figure 1. Screen Shot of Product Listing Page

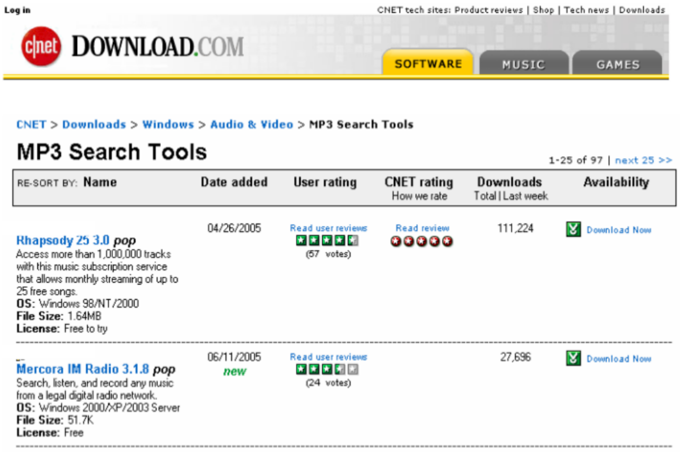
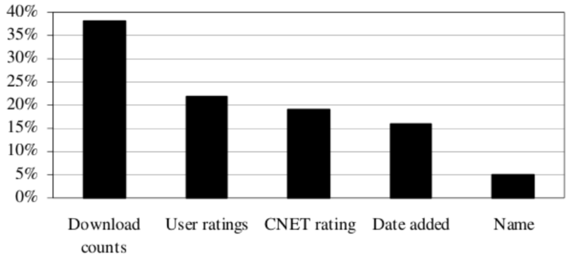
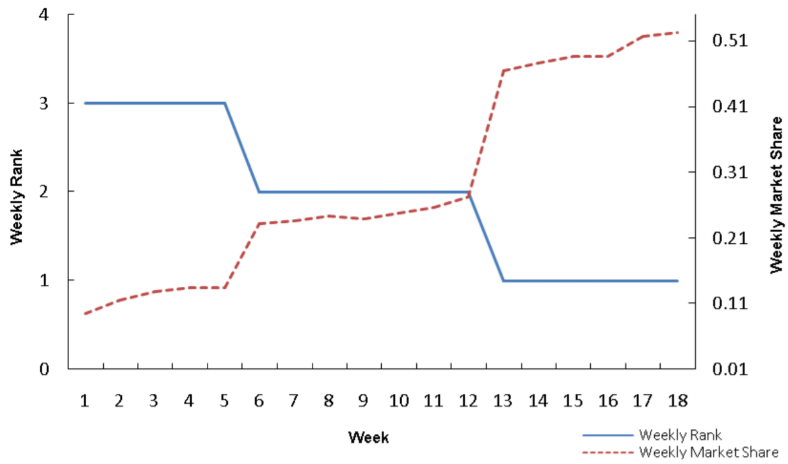


Figure 2. Popularity of Sort Options on Download.com Listing Pages



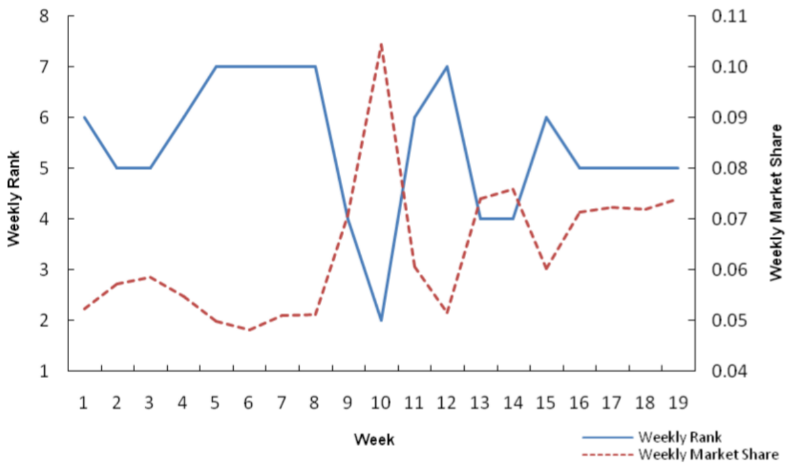
Adapted from Download.com report

Figure 3a. Weekly Rank and Weekly Market Share for Software "Bitcomet"



Note: "Bitcomet" is the software to share and download files from BitTorrent and chat with friends. This software is listed under "File Sharing" on CNED.

Figure 3b. Weekly Rank and Weekly Market Share for Software "Star Dowloader"



Note: "Star Downloader" is the software to Split files into several parts and download them simultaneously. This software is listed under "Download Managers" on CNED.

Table 1. Customer Type and Choice Set

Customer Type	Rank #1 (A)	Rank #2 (B)	Rank #3 (C)
{A, B, C}	X		
{A, B}	X		
{A, C}	X		
{B, C}		X	
{A}	X		
{B}		X	
{C}			X

Table 2. Summary of the Software Categories

Software Category	Daily Number of Software Programs Listed
Adware & Spyware Removal	90 – 138
Browsers	112 – 137
Download Managers	113 – 127
File Compression	59 – 83
File Sharing	102 – 132
Internet Chat	254 – 282
Media Players	154 – 185
MP3 Search Tools	73 – 123

Note: number of software is based on data from November 2004 to June 2005

Table 3. Description of Key Variables

Variable	Description and Measure
$TOTALDOWNLOAD_{it}$	Cumulative total number of downloads of software *i* at time *t* (in millions)
$LASTWEEKDOWNLOAD_{it}$	The most recent weekly number of downloads of software *i* at time *t* (in millions)
$WEEKLYRANK_{it}$	The rank of the most recent weekly number of downloads of software *i* at time *t*
$USERRATINGD_{iti}$	A dummy variable measures if software *i* has user rating at time *t*
$USERRATING_{it}$	Average user ratings for software *i* at *t* (one to five scale with half points)
$CNETRATINGD_{it}$	A dummy variable measures if software *i* has CNET rating at time *t*
$CNETRATING_{i}$	CNET rating for software *i* (one-to-five scale)
AGE_{it}	Number of days that software *i* has been listed up to time *t*
$POPD_{it}$	A dummy variable measures if software *i* is labeled as *popular* at time *t*
$NEWD_{it}$	A dummy variable measures if software *i* is labeled as *new* at time *t*
$NUMSOFTWARE_{t}$	Number of software listed in a specific category at time *t*

Table 4. Distribution of the Most Popular Software Categories

Most Popular Software Category	Software Count
MP3 Search Tools	14
Adware & Spyware Removal	13
Internet Chat	11
File Sharing	6
Antivirus	5
Browsers	5
Media Players	5
System Utilities	3
Others	44
Total	106

Note: number of software is based on data from November 2004 to June 2005

Table 5. Distributions of the CNET and User Ratings of Most Popular Software Programs

Rating	Frequency		
	User Rating Count	CNET Rating Count	User Rating Count With no CNET Rating
1-2	5 (5%)	1 (1%)	2 (5%)
2-3	7 (8%)	7 (10%)	2 (5%)
3-4	45 (49%)	26 (38%)	18 (48%)
4-5	35 (38%)	35 (51%)	8 (22%)
Total	92	69	37

Note: For both user and CNET ratings, the scale is from 1-5 with '1'=worst and '5' = best. There are 14 software programs that have not been rated by users, 37 not rated by CNET, and 7 that have neither user nor CNET ratings

Table 6a. Descriptive Statistics of Key Variables

Variable	N	Mean	Median	S. D.	Min.	Max.
Adware & Spyware Removal						
WEEKLYSHARE	1713	0.01	0.0001	0.06	0.00	0.75
WEEKLYRANK	1713	52.06	51.00	30.96	1.00	124.00
TOTALDOWNLOAD (million)	1713	2.22	0.04	13.60	45.00	136.00
LASTWEEKDOWNLOAD (million)	1713	0.03	0.0005	0.19	0.00	3.49
AGE	1713	147.21	88.00	168.12	0.00	847.00
CNETRATING	635	3.80	4.00	0.89	2.00	5.00
USERRATING	1538	3.62	3.50	1.22	1.00	5.50
POPD	1713	0.51	1.00	0.50	0.00	1.00
NEWD	1713	0.15	0.00	0.35	0.00	1.00
NUMSOFTWARE	1713	105.11	104.00	9.64	90.00	138.00
Browsers						
WEEKLYSHARE	2314	0.01	0.0001	0.04	0.00	0.56
WEEKLYRANK	2314	61.98	63.00	33.67	1.00	135.00
TOTALDOWNLOAD (million)	2314	0.36	0.007	1.45	10.00	11.40
LASTWEEKDOWNLOAD (million)	2314	0.002	0.00004	0.01	0.00	0.23
AGE	2314	600.22	632.00	498.82	0.00	3,067.00
CNETRATING	462	3.90	4.00	0.66	3.00	5.00
USERRATING	1857	3.66	3.50	1.17	1.00	5.00
POPD	2314	0.41	0.00	0.49	0.00	1.00
NEWD	2314	0.05	0.00	0.22	0.00	1.00
NUMSOFTWARE	2314	118.14	115.00	7.26	112.00	137.00
Download Managers						
WEEKLYSHARE	2164	0.01	0.001	0.03	0.00	0.38
WEEKLYRANK	2164	60.02	59.00	34.94	1.00	125.00
TOTALDOWNLOAD (million)	2164	0.76	0.02	4.81	17.00	48.80
LASTWEEKDOWNLOAD (million)	2164	0.002	0.0001	0.006	0.00	0.07
AGE	2164	485.63	363.50	423.83	0.00	1,809.00
CNETRATING	599	3.67	4.00	0.82	2.00	5.00
USERRATING	1589	3.50	3.50	1.17	1.00	5.00
POPD	2164	0.44	0.00	0.50	0.00	1.00
NEWD	2164	0.04	0.00	0.20	0.00	1.00
NUMSOFTWARE	2164	119.90	120.00	4.45	113.00	127.00
File Compression						
WEEKLYSHARE	1266	0.02	0.0002	0.09	0.00	0.68
WEEKLYRANK	1266	33.16	33.00	19.20	1.00	82.00
TOTALDOWNLOAD (million)	1266	3.14	0.04	18.70	32.00	146.00
LASTWEEKDOWNLOAD (million)	1266	0.01	0.0002	0.06	0.00	0.47
AGE	1244	437.40	310.50	391.75	1.00	1,632.00
CNETRATING	286	3.89	4.00	0.50	3.00	5.00
USERRATING	855	3.92	4.00	1.10	1.00	5.00
POPD	1266	0.41	0.00	0.49	0.00	1.00
NEWD	1266	0.02	0.00	0.15	0.00	1.00
NUMSOFTWARE	1266	65.64	62.00	7.43	59.00	83.00

Table 6b. Descriptive Statistics of Key Variables

Variable	N	Mean	Median	S. D.	Min.	Max.
File Sharing						
WEEKLYSHARE	2356	0.01	0.0002	0.04	0.00	0.53
WEEKLYRANK	2356	60.86	61.00	35.15	1.00	130.00
TOTALDOWNLOAD (million)	2356	0.25	0.01	0.85	6.00e-6	7.41
LASTWEEKDOWNLOAD (million)	2356	0.004	0.0001	0.02	0.00	0.33
AGE	2347	364.14	210.00	349.99	0.00	1,207.00
CNETRATING	228	3.73	4.00	0.84	1.00	5.00
USERRATING	1869	3.48	3.50	1.23	1.00	5.00
POPD	2356	0.42	0.00	0.49	0.00	1.00
NEWD	2356	0.06	0.00	0.23	0.00	1.00
NUMSOFTWARE	2356	119.47	121.00	8.83	102.00	132.00
Internet Chat						
WEEKLYSHARE	4769	0.004	0.00004	0.04	0.00	0.67
WEEKLYRANK	4769	133.92	135.50	76.36	1.00	280.00
TOTALDOWNLOAD (million)	4769	1.78	0.01	16.91	0.00005	255.00
LASTWEEKDOWNLOAD (million)	4769	0.005	0.00006	0.05	0.00	1.30
AGE	4769	611.31	534.00	514.23	0.00	2,666.00
CNETRATING	300	3.83	4.00	0.84	2.00	5.00
USERRATING	4120	3.99	4.50	1.01	1.00	5.00
POPD	4769	0.12	0.00	0.33	0.00	1.00
NEWD	4769	0.05	0.00	0.21	0.00	1.00
NUMSOFTWARE	4769	264.90	264.00	6.73	254.00	282.00
Media Player						
WEEKLYSHARE	3181	0.005	0.0004	0.02	0.00	0.33
WEEKLYRANK	3181	87.75	88.00	50.26	1.00	184.00
TOTALDOWNLOAD (million)	3181	0.59	0.04	2.69	0.00005	33.80
LASTWEEKDOWNLOAD (million)	3181	0.003	0.0003	0.02	0.00	0.22
AGE	3181	439.96	269.00	464.40	0.00	2,477.00
CNETRATING	614	3.53	4.00	1.05	1.00	5.00
USERRATING	2620	3.60	3.50	1.27	1.00	5.00
POPD	3181	0.17	0.00	0.38	0.00	1.00
NEWD	3181	0.09	0.00	0.29	0.00	1.00
NUMSOFTWARE	3181	174.23	176.00	8.29	154.00	185.00
MP3 Search Tools						
WEEKLYSHARE	1981	0.01	0.0004	0.05	0.00	0.54
WEEKLYRANK	1981	53.51	51.00	32.35	1.00	122.00
TOTALDOWNLOAD (million)	1981	3.18	0.07	15.20	0.0001	134.00
LASTWEEKDOWNLOAD (million)	1981	0.03	0.001	0.14	0.00	1.52
AGE	1981	267.27	117.00	325.99	0.00	1,652.00
CNETRATING	259	3.52	4.00	0.77	2.00	5.00
USERRATING	1914	3.36	3.50	0.99	1.00	5
POPD	1981	0.28	0.00	0.45	0.00	1.00
NEWD	1981	0.18	0.00	0.38	0.00	1.00
NUMSOFTWARE	1981	109.16	117.00	13.64	73.00	123.00

Table 7a. Correlation Matrix of Key Variables

Variable	1	2	3	4	5	6	7	8	9	10
Adware & Spyware Removal										
1 lg$WEEKLYSHARE_{it}$	1									
2 $WEEKLYRANK_{i,t-1}$	-0.92	1								
3 $TOTALDOWNLOAD_{i,t-1}$	0.46	-0.26	1							
4 $LASTWEEKDOWNLOAD_{it}$	0.48	-0.27	0.68	1						
5 $POPD_{it}$	0.77	-0.72	0.16	0.17	1					
6 $NEWD_{it}$	0.17	-0.14	-0.004	-0.01	0.14	1				
7 $CNETRATINGD_{i,t-1}$	0.39	-0.36	0.20	0.21	0.24	0.006	1			
8 $USERRATING_{i,t-1}$	0.28	-0.27	0.11	0.12	0.25	-0.07	0.14	1		
9 AGE_{it}	-0.15	0.19	0.05	0.03	-0.17	-0.34	0.04	0.06	1	
10 $NUMSOFTWARE_{it}$	-0.09	0.11	-0.004	-0.02	-0.06	-0.06	-0.02	-0.04	0.04	1
Browsers										
1 lg$WEEKLYSHARE_{it}$	1									
2 $WEEKLYRANK_{i,t-1}$	-0.90	1								
3 $TOTALDOWNLOAD_{i,t-1}$	0.61	-0.40	1							
4 $LASTWEEKDOWNLOAD_{it}$	0.53	-0.31	0.67	1						
5 $POPD_{it}$	0.75	-0.72	0.29	0.21	1					
6 $NEWD_{it}$	0.29	-0.20	0.21	0.31	0.19	1				
7 $CNETRATINGD_{i,t-1}$	0.43	-0.39	0.18	0.27	0.41	0.14	1			
8 $USERRATING_{i,t-1}$	0.35	-0.39	0.11	0.12	0.31	0.04	0.02	1		
9 AGE_{it}	-0.16	0.13	-0.03	-0.16	-0.17	-0.27	-0.37	0.11	1	
10 $NUMSOFTWARE_{it}$	-0.02	0.11	0.005	-0.02	-0.07	-0.05	0.008	-0.002	0.03	1
Download Managers										
1 lg$WEEKLYSHARE_{it}$	1									
2 $WEEKLYRANK_{i,t-1}$	-0.93	1								
3 $TOTALDOWNLOAD_{i,t-1}$	0.38	-0.24	1							
4 $LASTWEEKDOWNLOAD_{it}$	0.57	-0.40	0.65	1						
5 $POPD_{it}$	0.68	-0.73	0.16	0.28	1					
6 $NEWD_{it}$	0.18	-0.09	0.02	0.09	0.10	1				
7 $CNETRATINGD_{i,t-1}$	0.18	-0.13	0.23	0.23	0.13	0.08	1			
8 $USERRATING_{i,t-1}$	0.32	-0.35	0.02	0.08	0.24	-0.02	0.01	1		
9 AGE_{it}	-0.42	0.41	-0.09	-0.22	-0.32	-0.24	-0.31	0.17	1	
10 $NUMSOFTWARE_{it}$	-0.04	0.04	0.01	0.01	-0.03	-0.001	0.01	-0.001	-0.02	1
File Compression										
1 lg$WEEKLYSHARE_{it}$	1									
2 $WEEKLYRANK_{i,t-1}$	-0.93	1								
3 $TOTALDOWNLOAD_{i,t-1}$	0.51	-0.28	1							
4 $LASTWEEKDOWNLOAD_{it}$	0.56	-0.31	0.68	1						
5 $POPD_{it}$	0.74	-0.68	0.20	0.23	1					
6 $NEWD_{it}$	0.06	-0.04	0.03	0.02	0.05	1				
7 $CNETRATINGD_{i,t-1}$	0.36	-0.28	0.30	0.32	0.22	0.003	1			
8 $USERRATING_{i,t-1}$	0.23	-0.24	0.08	0.10	0.27	-0.03	0.06	1		
9 AGE_{it}	-0.11	0.07	-0.10	-0.12	-0.15	-0.17	-0.10	0.39	1	
10 $NUMSOFTWARE_{it}$	-0.05	0.14	-0.001	-0.01	-0.03	-0.05	0.05	-0.01	0.05	1

Table 7b. Correlation Matrix of Key Variables

Variable	1	2	3	4	5	6	7	8	9	10
File Sharing										
1 lg$WEEKLYSHARE_{it}$	1									
2 $WEEKLYRANK_{i,t-1}$	-0.93	1								
3 $TOTALDOWNLOAD_{i,t-1}$	0.52	-0.40	1							
4 $LASTWEEKDOWNLOAD_{it}$	0.46	-0.31	0.73	1						
5 $POPD_{it}$	0.68	-0.73	0.32	0.22	1					
6 $NEWD_{it}$	0.17	-0.11	0.10	0.10	0.12	1				
7 $CNETRATINGD_{i,t-1}$	0.23	-0.17	0.30	0.37	0.19	0.02	1			
8 $USERRATING_{i,t-1}$	0.27	-0.27	0.16	0.12	0.20	0.04	0.11	1		
9 AGE_{it}	-0.33	0.31	-0.10	-0.14	-0.29	-0.25	-0.11	0.14	1	
10 $NUMSOFTWARE_{it}$	-0.07	0.09	0.03	0.03	-0.06	-0.004	0.01	0.01	0.006	1
Internet Chat										
1 lg$WEEKLYSHARE_{it}$	1									
2 $WEEKLYRANK_{i,t-1}$	-0.89	1								
3 $TOTALDOWNLOAD_{i,t-1}$	0.35	-0.18	1							
4 $LASTWEEKDOWNLOAD_{it}$	0.39	-0.18	0.45	1						
5 $POPD_{it}$	0.68	-0.53	0.26	0.27	1					
6 $NEWD_{it}$	0.19	-0.15	-0.01	0.02	0.12	1				
7 $CNETRATINGD_{i,t-1}$	0.57	-0.35	0.38	0.38	0.56	0.08	1			
8 $USERRATING_{i,t-1}$	0.12	-0.16	0.01	0.04	0.03	-0.02	0.01	1		
9 AGE_{it}	-0.14	0.09	0.001	-0.09	-0.16	-0.26	-0.18	0.20	1	
10 $NUMSOFTWARE_{it}$	-0.02	0.03	0.001	-0.001	0.003	-0.02	0.001	-0.002	0.02	1
Media Player										
1 lg$WEEKLYSHARE_{it}$	1									
2 $WEEKLYRANK_{i,t-1}$	-0.94	1								
3 $TOTALDOWNLOAD_{i,t-1}$	0.44	-0.32	1							
4 $LASTWEEKDOWNLOAD_{it}$	0.45	-0.31	0.68	1						
5 $POPD_{it}$	0.62	-0.55	0.34	0.38	1					
6 $NEWD_{it}$	0.20	-0.14	-0.05	-0.03	0.11	1				
7 $CNETRATINGD_{i,t-1}$	0.31	-0.28	0.32	0.25	0.19	-0.07	1			
8 $USERRATING_{i,t-1}$	0.17	-0.17	0.08	0.05	0.07	0.003	0.06	1		
9 AGE_{it}	-0.09	0.08	0.08	-0.03	-0.07	-0.29	-0.05	0.25	1	
10 $NUMSOFTWARE_{it}$	0.001	0.05	-0.004	0.001	-0.002	0.005	0.009	-0.001	-0.01	1
MP3 Search Tools										
1 lg$WEEKLYSHARE_{it}$	1									
2 $WEEKLYRANK_{i,t-1}$	-0.91	1								
3 $TOTALDOWNLOAD_{i,t-1}$	0.47	-0.30	1							
4 $LASTWEEKDOWNLOAD_{it}$	0.50	-0.30	0.57	1						
5 $POPD_{it}$	0.70	-0.64	0.30	0.30	1					
6 $NEWD_{it}$	0.26	-0.23	0.09	0.03	0.12	1				
7 $CNETRATINGD_{i,t-1}$	0.49	-0.39	0.45	0.44	0.38	0.01	1			
8 $USERRATING_{i,t-1}$	0.12	-0.11	0.01	0.11	0.11	0.09	0.13	1		
9 AGE_{it}	-0.42	0.44	-0.12	-0.12	-0.14	-0.37	-0.10	-0.05	1	
10 $NUMSOFTWARE_{it}$	-0.08	0.19	-0.03	-0.02	-0.14	-0.03	-0.05	-0.08	0.04	1

Table 8. Fixed Effects Estimation Results

	Adware & Spyware Removal	Browsers	Download Managers	File Compression	File Sharing	Internet Chat	Media Players	MP3 Search Tools
Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
WEEKLYRANK_1$_{i,t-1}$	-0.02***	-0.02***	-0.01***	-0.02***	-0.02***	-0.01***	-0.01***	-0.01***
	(0.002)	(0.002)	(0.001)	(0.004)	(0.002)	(0.001)	(0.001)	(0.003)
WEEKLYRANK_1SQ$_{i,t-1}$	0.00004**	0.0001***	0.00004***	0.00003	0.0001***	0.00003***	0.00003***	0.00004**
	(0.00002)	(0.00001)	(0.00001)	(0.00004)	(0.00001)	(0.000001)	(0.000004)	(0.000001)
USERRATING$_{i,t-1}$	-0.03	-0.04	0.01	-0.04	-0.001	-0.004	-0.02	-0.02
	(0.02)	(0.03)	(0.01)	(0.03)	(0.02)	(0.02)	(0.01)	(0.03)
*WEEKLYRANK_1*USERRATING*	0.001***	0.001***	0.0004**	0.002***	0.001***	0.0002***	0.001***	0.001***
	(0.0004)	(0.0003)	(0.0002)	(0.0005)	(0.0002)	(0.00007)	(0.0001)	(0.0004)
USERRATINGD$_{i,t-1}$	-0.09	0.14**	-0.19***	0.05	-0.13	-0.11	-0.02	0.09
	(0.09)	(0.07)	(0.06)	(0.09)	(0.09)	(0.09)	(0.05)	(0.20)
CNETRATINGD$_{i,t-1}$	-0.03	0.04	0.09***	-0.07	-0.13	0.06	0.28***	0.03
	(0.07)	(0.05)	(0.03)	(0.04)	(0.12)	(0.06)	(0.07)	(0.14)
LASTWEEKDOWNLOAD$_{it}$	0.76**	1.46**	9.64***	2.90*	1.02	0.93***	0.15	0.35
	(0.31)	(0.68)	(3.09)	(1.70)	(0.79)	(0.19)	(2.24)	(0.21)
TOTALDOWNLOAD$_{i,t-1}$	0.01	0.09	0.38	0.02	0.02	0.32***	-0.07	0.02
	(0.01)	(0.06)	(0.36)	(0.01)	(0.05)	(0.06)	(0.70)	(0.02)
POPD$_{it}$	0.48***	0.14***	0.13***	0.11***	0.15***	0.20***	0.39***	0.29***
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
NEWD$_{it}$	0.05*	0.11***	0.05**	0.04	0.01	0.18***	0.18***	0.14***
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
AGESQ$_{i,t-1}$	2.65e-06	8.96e-07	3.45e-06***	1.32e-06	2.03e-06	6.01e-06	2.89e-06	-1.64e-06
	(0.0004)	(6.10e-07)	(1.29e-06)	(1.00e-06)	(2.81e-06)	(1.68e-06)	(2.93e-06)	(2.46e-06)
NUMSOFTWARE$_{it}$	-0.01***	-0.01**	-0.01***	0.01***	-0.007***	0.001	-0.001*	-0.001
	(0.004)	(0.003)	(0.002)	(0.004)	(0.001)	(0.001)	(0.0007)	(0.001)
AGE$_{it}$	*coefficients on AGE are estimated for each individual product*							
	N = 1479	*N = 2085*	*N = 1903*	*N = 1123*	*N = 2078*	*N = 4218*	*N = 2794*	*N = 1699*
	Group = 98	*Group = 110*	*Group = 125*	*Group = 59*	*Group = 126*	*Group = 263*	*Group = 191*	*Group = 133*
	R^2 = 0.50	*R^2 = 0.43*	*R^2 = 0.45*	*R^2 = 0.34*	*R^2 = 0.42*	*R^2 = 0.55*	*R^2 = 0.30*	*R^2 = 0.28*

*Note: *** p<.01 ** p<.05 * p<.10 standard errors in parentheses*

Software dummies (fixed effects for each software) and time dummies (fixed effects for each week) used in estimating the model are not reported.